Airy Phenomena and Analytic Combinatorics of Connected Graphs

Philippe Flajolet, Bruno Salvy, Gilles Schaeffer

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Airy Phenomena and Analytic Combinatorics of Connected Graphs

Philippe FLAJOLET, Bruno SALVY, Gilles SCHAEFFER

N° 4581

Octobre 2002

THÈME 2

apport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4581--FR+ENG

Airy Phenomena and Analytic Combinatorics of Connected Graphs

Philippe FLAJOLET, Bruno SALVY, Gilles SCHAEFFER

Thème 2 — Génie logiciel
et calcul symbolique
Projet Algorithmes

Rapport de recherche n° 4581 — Octobre 2002 — 25 pages

Abstract: Until now, the enumeration of connected graphs has been dealt with by probabilistic methods, by special combinatorial decompositions or by somewhat indirect formal series manipulations. We show here that it is possible to make analytic sense of the divergent series that expresses the generating function of connected graphs. As a consequence, it becomes possible to derive analytically known enumeration results using only first principles of combinatorial analysis and straight asymptotic analysis—specifically, the saddle-point method. In this perspective, the enumeration of connected graphs by excess (of number of edges over number of vertices) derives from a simple saddle-point analysis. Furthermore, a refined analysis based on coalescent saddle points yields complete asymptotic expansions for the number of graphs of fixed excess, through an explicit connection with Airy functions.

Key-words: analytic combinatorics, graph theory, random graphs, connectivity, Airy function

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Phénomène d'Airy et
Combinatoire Analytique des Graphes Connexes

Résumé : Jusqu'à présent, le dénombrement des graphes connexes a été traité par des méthodes probabilistes, par des décompositions combinatoires particulières ou par des manipulations de séries assez indirectes. Nous montrons ici qu'il est possible de donner un sens analytique à la série divergente qui exprime la fonction génératrice des graphes connexes. En conséquence, il devient possible d'obtenir analytiquement des résultats de dénombrement connus en utilisant uniquement des principes élémentaires d'analyse combinatoire et de l'analyse asymptotique classique (la méthode du col). Dans cette perspective, le dénombrement des graphes connexes par excès (du nombre d'arêtes sur le nombre de sommets) se déduit d'une analyse de col simple. De plus, un raffinement de l'analyse fondée sur des points cols coalescents donne un développement asymptotique complet pour le nombre de graphes d'excès fixé, via une connection explicite avec les fonctions d'Airy.

Mots-clés : combinatoire analytique, théorie des graphes, graphes aléatoires, connectivité, fonction d'Airy

AIRY PHENOMENA AND ANALYTIC COMBINATORICS OF CONNECTED GRAPHS

PHILIPPE FLAJOLET, BRUNO SALVY, AND GILLES SCHAEFFER

ABSTRACT. Until now, the enumeration of connected graphs has been dealt with by probabilistic methods, by special combinatorial decompositions or by somewhat indirect formal series manipulations. We show here that it is possible to make analytic sense of the divergent series that expresses the generating function of connected graphs. As a consequence, it becomes possible to derive analytically known enumeration results using only first principles of combinatorial analysis and straight asymptotic analysis—specifically, the saddle-point method. In this perspective, the enumeration of connected graphs by excess (of number of edges over number of vertices) derives from a simple saddle-point analysis. Furthermore, a refined analysis based on coalescent saddle points yields complete asymptotic expansions for the number of graphs of fixed excess, through an explicit connection with Airy functions.

INTRODUCTION

E. M. Wright, of Hardy and Wright fame, initiated the enumeration of labelled connected graphs by number of vertices and edges in a well-known series of articles [34, 35, 36]. In particular, he discovered that the generating functions of graphs with a fixed excess of number of edges over number of vertices has a rational expression in terms of the tree function $T(z)$. Wright's approach is based on the fact that deletion of an edge in a connected graph leads to either one or two connected graphs with smaller excess. This decomposition translates into a quadratic differential recurrence from which Wright was able to deduce general structural results, especially as regards dominant asymptotics.

The problem of enumerating connected graphs by excess is obviously related to the question of connectivity in random graphs and so, not unnaturally, it has been also approached repeatedly through the probabilistic method. It is for instance of special importance in the emergence of the "giant component" under Erdős and Rényi's model. Bollobás's book [5, Ch. 6] contains an account of various aspects of the question examined from the probabilistic angle. The "giant paper" of Janson, Knuth, Łuczak and Pittel [18] devotes some 25 pages to generating function evaluations before going into the actual physics of the random graph phase transition. Finally, the enumerative results valid asymptotically over the widest range of the parameters are those of Bender, Canfield, and McKay in [3].

In contrast, our approach here is completely straightforward. It starts from the bivariate generating function of connected graphs

$$C(z,q) = \log\left(1 + \sum_{n\geq 1}(1+q)^{n(n-1)/2}\frac{z^n}{n!}\right), \tag{1}$$

Date: October 11, 2002.

that is viewed nowadays as a direct instance of the classical "exponential formula" of combinatorial analysis; see for instance [15, 32]. (The formula was published by Riddel and Uhlenbeck [26] in 1953.) We show that this series that strongly diverges for any $q > 0$ can in fact be represented by an integral that gives it *bona fide* analytic meaning for small $q < 0$. In a way, this amounts to assigning *negative* weights (or probabilities) to edges, contrary to what is done commonly in probabilistic or enumerative treatments of the question like [3, 5]. Then, standard methods of asymptotic analysis, especially the saddle-point technique, apply. Thus, in a logical sense, the enumeration of graphs by excess "only" requires the exponential formula and basic asymptotics. Together with the article of Janson *et al.* [18], the present article is one of the very few approaches that treats connectivity of graphs starting from first principles. As opposed to [18], our approach is purely analytic and hopefully a little more transparent from a logical standpoint. It is also a curious fact that asymptotic analysis is used here to establish an exact enumerative result.

Our principal result is a purely analytic proof of a theorem, known from earlier works of Wright[1] and of Janson *et al.* [18, 34]. A main character throughout the article is the "tree function" that is defined by

$$T(z) = ze^{T(z)}, \qquad T(z) = \sum_{n\geq 1} n^{n-1}\frac{z^n}{n!}, \tag{2}$$

and is otherwise known to enumerate rooted labelled trees. For any connected graph with k edges and n vertices, the quantity $k-n$ is always at least -1 and is called the *excess*[2] of the graph. Our goal is a characterization of the (exponential) generating functions (GFs) of graphs of any fixed excess.

Theorem 1. (*i*) *The generating function of unrooted trees (graphs with excess* -1*) is*

$$W_{-1}(z) = T(z) - \frac{1}{2}T^2(z). \tag{3}$$

(*ii*) *The GF of connected graphs with excess* 0 *(unicyclic graphs) is*

$$W_0(z) = \frac{1}{2}\log\frac{1}{1-T(z)} - \frac{1}{2}T(z) - \frac{1}{4}T^2(z). \tag{4}$$

(*iii*) *The GF of connected graphs with excess* $k \geq 1$ *is a rational function of* $T(z)$*: there exist polynomials* A_k*, such that*

$$W_k(z) = \frac{A_k(T(z))}{(1-T(z))^{3k}}. \tag{5}$$

Part (*i*) is commonly attributed to Cayley and several of his contemporaries (see [18, p. 240] for a discussion), while Part (*ii*) is due to Rényi; Equation (5) of Part (*iii*) is Wright's main result. Observe that Wright had to resort to an "external argument" based on special multigraphs [34, Sec. 7] in order to obtain the rationality of the $W_k(z)$ in terms of $T(z)$.

For completeness, we recall that the generating functions provided by Theorem 1 are equivalent to explicit forms for the quantities[3] $C_{n,n+\ell} = n![z^n]W_\ell(z)$, as results

[1]Wright's results were to some extent anticipated by Temperley [31] whose insightful short note of 1959 seems to rely partly on heuristic arguments.

[2]Our notion of excess is consistent with the one of Janson *et al.* [18, p. 240]. Our W_k coincide with those of Wright [34, p. 318] and are equal to the $\widehat{C}_k$ in the notations of [18].

[3]As usual, we denote by $[z^n]f(z)$ the nth coefficient in the series $f(z)$.

from the standard expansions,

$$\frac{1}{1-T(z)} = 1 + \sum_{n\geq 1} n^n \frac{z^n}{n!}, \qquad \log \frac{1}{1-T(z)} = \sum_{n\geq 0} Q_n n^{n-1} \frac{z^n}{n!}, \tag{6}$$

where Q_n is the Ramanujan Q-function (see [19, 9] and references therein):

$$Q_n = 1 + \frac{n-1}{n} + \frac{(n-1)(n-2)}{n^2} + \cdots .$$

The numerical coefficients $A_k(1)$ are crucial to dominant asymptotics. Indeed, from either the explicit forms deriving from (6) or from the known singular expansion [21] of the tree function, namely $T(z) = 1 - \sqrt{2}\sqrt{1-ez} + \cdots$, the following holds.

Corollary 1. *The asymptotic form of the graph counts* $C_{n,n+k} = n![z^n]W_k(z)$ *for fixed* $k \geq 2$ *is*

$$A_k(1)\sqrt{\pi}\left(\frac{n}{e}\right)^n \left(\frac{n}{2}\right)^{\frac{3k-1}{2}} \left(\frac{1}{\Gamma\left(\frac{3k}{2}\right)} + \frac{\frac{A_k'(1)}{A_k(1)} - k}{\Gamma\left(\frac{3k-1}{2}\right)}\sqrt{\frac{2}{n}} + O\left(\frac{1}{n}\right)\right), \tag{7}$$

and more generally, lower order terms depend on the derivatives $A_k^{(j)}(1)$.

Our analysis allows us to characterize these coefficients.

Theorem 2. (i) *The generating function of the dominant coefficients* $A_k(1)$ *is expressible as*

$$\sum_{k=1}^{\infty} A_k(1)x^k = \log\left(\sum_{k=0}^{\infty} c_k x^k\right), \tag{8}$$

where

$$c_k = \frac{(-1)^k\Gamma(3k+\frac{1}{2})}{54^k k!\Gamma(k+\frac{1}{2})} = \frac{(-1)^k(6k)!}{(3k)!(2k)!3^{2k}2^{5k}}.$$

(ii) *For* $j \geq 1$, *the generating function of the derivatives* $A_k^{(j)}(1)$ *can be expressed in terms of the classical Airy function*

$$\mathrm{Ai}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(zt+t^3/3)}dt, \quad \textit{solution of } y'' + zy = 0.$$

More precisely, let

$$S(x) = -\frac{2}{x}\left(1 + (2x)^{1/3}\frac{\mathrm{Ai}'((2x)^{-2/3})}{\mathrm{Ai}((2x)^{-2/3})}\right) = 1 + \frac{95}{288}x + O(x^2), \quad x \to 0,$$

then

$$\sum_{k\geq 1} A_k^{(j)}(1)x^k = \mathcal{A}^{(j)}(x, S(x)),$$

where $\mathcal{A}^{(j)}(x,s)$ *is a polynomial of degree* j *in* s *with coefficients that are Laurent polynomials in* x. *These polynomials can be determined effectively from Equation (41) below. (See Appendix II for a table.)*

The coefficients $A_k^{(j)}(1)$ intervene as subdominant terms in the expansion of $W_k(z)$ and their knowledge provides a full asymptotic expansion extending (7).

Part (i), the form (8) of the driving coefficients $A_k(1)$, is given explicitly by Janson *et al.* in [18] where the authors built upon earlier results of the "Russian school", most notably Bagaev and Voblyi. In view of Corollary 1 and the accompanying remarks, this form characterizes the dominant asymptotics of the number of graphs of some fixed excess k. Part (ii) of Theorem 2 then provides a "correction series" that describes precisely what goes on in successive subdominant asymptotic terms.

Wright had in fact obtained earlier a recursive determination of the $A_k(1)$ but he does not appear to have obtained the relation (8). We now know that the $W_k(z)$ and the $A_k(1)$ intervene in a number of closely interrelated problems and a variant of the sequence $\{A_k(1)\}$, called the "Wright-Louchard-Takács sequence", appears in [13]. Indeed, the $W_k(z)$ and the $A_k(1)$ surface in such diverse problems as: parking and linear probing hashing [13, 20], Brownian excursion area and area below Dyck path [22, 11], area below the Poisson excursion [27], inversions in trees [14, 23], path length in trees of various sorts [30, 28, 29], and naturally the enumeration of connected graphs [10, 18, 34, 35, 36]. See [13, 20] for a combinatorial perspective on the relationship between these problems.

The present article fits in a more global endeavour to find simple reasons for the occurrence of the Airy function in so many problems of analytic combinatorics. One good reason is the connection with coalescing saddles as exemplified by [2] in the case of random maps that are random planar graphs of a specific type. We propose to examine in future works the extent to which it can be applied to other graph models, to uniform estimates, and to phase transitions that arise in hashing and random allocation problems [13, 20].

Finally, statistical physics is lurking in the background. In a partly heuristic, but insightful paper [31], Temperley developed formulæ that correspond to a primitive form of Theorem 1. In [24], Monasson proposed to approach connectivity of the random graph via the replica method. It would be of obvious interest to confront the rigourous approach developed here with the powerful (but yet unrigourous) replica method. The present paper may hopefully contribute to the debate.

Plan of the article. The article is entirely based on an integral representation for the divergent series in (1). In other words, the generating function of graphs can be viewed as the asymptotic expansion of a *bona fide* analytic object. This is described in Section 1, where a combinatorial bijection due to Gessel and Wang is used to dispose of some of the divergence issues. A straightforward application of the saddle-point method for the asymptotics of integrals then yields Equations (3), (4) and (5) in Section 2. Consequently, Wright's representation (Part (iii) of Theorem 1) appears to *coincide* with a standard saddle-point expansion. The dominant coefficients $A_k(1)$ as well as their their subdominant companions, $A'_k(1)$, etc., are a bit more recondite. In Section 3, we show that relevant information can be gathered by a method of coalescent saddle points, which gives Theorem 2. Thus, it appears that very detailed formal expansions found by Janson *et al.* [18] are in fact precisely double saddle-point expansions. The expressions obtained involve hypergeometric functions that are reducible to the classical Airy function.

1. Principles of an Analytic Approach

We discuss here the principles on which our proof of Theorem 1 is built. It is based on assigning a complex-analytic meaning to strongly divergent series that express graphical enumerations and to correlative series rearrangements. Accordingly, special attention is required in distinguishing carefully between formal objects and their analytic counterparts.

1.1. **Formal Expressions.** In what follows, for K a field and $z_1, z_2, \dots$ a collection of indeterminates, we let $K[[z_1, z_2, \dots]]$ denote the ring of formal power series with indeterminates $z_1, z_2, \dots$ and coefficients in K.

Let $G_{n,k}$ be the number of labelled graphs with n vertices and k edges and $C_{n,k}$ the number of those that are connected. The (formal) generating functions are defined as objects of $\mathbb{C}[[z,q]]$ by

$$G(z,q) := \sum_{n,k} G_{n,k} q^k \frac{z^n}{n!}, \qquad C(z,q) := \sum_{n,k} C_{n,k} q^k \frac{z^n}{n!}.$$

A graph is determined by the selection of edges amongst all possible pairs of points, implying the identity in $\mathbb{C}[[z,q]]$,

$$G(z,q) = \sum_{n\geq 0} (1+q)^{n(n-1)/2} \frac{z^n}{n!}. \tag{9}$$

On the other hand, a graph is a set of connected components, which, by the classical exponential formula, implies the relation $G(z,q) = \exp(C(z,q))$ and consequently

$$\begin{aligned} C(z,q) &= \log(G(z,q)) \\ &= z + q\frac{z^2}{2!} + (3q^2+q^3)\frac{z^3}{3!} + (16q^3 + 15q^4 + 6q^5 + q^6)\frac{z^4}{4!} + \cdots, \end{aligned} \tag{10}$$

valid again in $\mathbb{C}[[z,q]]$.

Consider next the GF of connected graphs counted according to excess and to size. An essential component of our approach is to record excess and do so by a *negatively signed* variable. Then, in $\mathbb{C}[[z,q]]$ we have

$$\begin{aligned} Q(z,q) &:= \sum_{n,\ell} C_{n,n+\ell}(-q)^{\ell+1}\frac{z^n}{n!} = -qC(-z/q,-q) \\ &= z + \frac{z^2}{2!} + (3-q)\frac{z^3}{3!} + (16 - 15q + 6q^2 - q^3)\frac{z^4}{4!} + \cdots \\ &= W_{-1}(z) - qW_0(z) + q^2W_1(z) - \cdots, \end{aligned}$$

where each $W_\ell(z) \in \mathbb{C}[[z]]$ is by definition the generating function of connected graphs with excess ℓ:

$$W_\ell(z) := \sum_n C_{n,n+\ell}\frac{z^n}{n!}. \tag{11}$$

For instance $W_{-1}(z)$ is the GF of unrooted trees, $W_0(z)$ the GF of unicyclic components, and so on.

Now, the exponential formula (10) in conjunction with (9) permits us to express $Q(z,q)$ formally as

$$Q(z,q) = -q\log\left(\sum_{n\geq 0}(1-q)^{n(n-1)/2}\frac{(-zq^{-1})^n}{n!}\right). \tag{12}$$

In this formula, the right-hand side is to be taken as an element of $\mathbb{C}(q)[[z]]$ (that is, the ring of formal power series in z, with coefficients that are rational functions in q) and the reorganization of the series takes place in that domain, according to the formula

$$\log(1+u) = \frac{u}{1} - \frac{u^2}{2} + \frac{u^3}{3} - \cdots.$$

Massive algebraic cancellations in the coefficient field $\mathbb{C}(q)$ take place when the series is reorganized and this seems to be the cause of many analytic hardships.

1.2. **Integral Representations.** The basic analytic representation derives from the following simple lemma that we state in its bare-bones version.

Lemma 1. *Let v_n be a finite sequence with generating function $V(z) = \sum_n v_n z^n$ and let w be a real number with $w \in (0,1)$. Then,*

$$\sum_n w^{n^2/2} v_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} V(e^{ix\sqrt{\log w^{-1}}}) e^{-x^2/2}\, dx. \tag{13}$$

The lemma directly results from the classical Fourier integral

$$e^{-t^2/2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ixt} e^{-x^2/2}\, dx.$$

The importance of (13) for the analysis of q-series comes from the fact that *the integral representation linearizes the quadratic forms present in the exponents.* Obviously, the lemma generalizes to infinite sequences that do not grow too fast. (See for instance [12] for a combinatorial application to chord systems.)

The graph generating function $G(z,q)$ specified by (9) and viewed as a function of its two parameters z,q diverges wildly as soon as q is positive. However, it acquires a *bona fide* analytic meaning if it is considered as a series in z with q a fixed parameter, provided $|1+q| < 1$. In that case, it becomes an entire function of z. Given this, we may legitimately expect $Q(z,q)$ to make analytic sense when q is restricted to the disk centred at $+1$ with radius 1. Precisely, we fix q as a *numerical* parameter such that $|1-q| < 1$ and consider the weighting π that assigns to a graph g the weight $\pi(g) := (-q)^{e(g)-|g|}$ where $e(g)$ is the number of edges and $|g|$ is the number of vertices of g. We introduce the two analytic objects

$$\mathcal{H}(z,q) := \sum_{g \text{ graph}} \pi(g) \frac{z^{|g|}}{|g|!}, \qquad \mathcal{Q}(z,q) := \sum_{g \text{ connected graph}} \pi(g) \frac{z^{|g|}}{|g|!}.$$

The function $\mathcal{H}$ is an entire function of z for q in the given range, since it is directly related to G by $\mathcal{H}(z,q) = G(-z/q, -q)$. The exponential connection between $\mathcal{H}$ and $\mathcal{Q}$, namely $\mathcal{Q} = \log \mathcal{H}$, holds. Observe also that $\mathcal{Q}$ is an analytic function of z for $|z|$ sufficiently small, since $\mathcal{H}(0,q) = 1$.

Application of Lemma 1 now yields the following basic integral representation.

Lemma 2. *The generating function of connected graphs counted by excess and weighted with negative weights admits for $q \in (0,1)$ the integral representation*

$$\mathcal{Q}(z,q) = -q \log \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2} - z \frac{(1-q)^{-1/2}}{q} e^{ix\lambda(q)} \right) dx \right), \tag{14}$$

where

$$\lambda \equiv \lambda(q) := \sqrt{\log(1-q)^{-1}} = \sqrt{q}\left(1 + \frac{1}{4}q + \frac{13}{96}q^2 + \cdots\right).$$

This representation is central to our treatment.

1.3. **Interchange of Limits and Coefficients.** We will prove later, as a by-product of the analysis, that there exists a family $\mathcal{W}_\ell(z)$ of analytic functions each having radius of convergence e^{-1} such that the analytic $\mathcal{Q}(z,q)$ satisfies when $|z| < e^{-1}$ as $q \to 0^+$:

$$\mathcal{Q}(z,q) \underset{q\to 0^+}{\sim} \sum_{\ell \geq 0} \mathcal{W}_{\ell-1}(z)(-q)^\ell. \tag{15}$$

The $\mathcal{W}_\ell$ found in the process satisfy precisely the conditions of Theorem 1 and the following sections will show how to establish (15) from first principles.

If granted the asymptotic expansion (15), the proof is complete once we establish that $W_\ell(z) = \mathcal{W}_\ell(z)$. Now, the *algebraic* quantities $W_\ell(z)$ are such that, by definition,

$$n![z^n]W_\ell(z) = C_{n,n+\ell} = (-1)^{\ell+1}[q^{\ell+1}]\left(n![z^n]Q(z,q)\right),$$

with $Q(z,q)$ as specified by the formal relation (12), and $[z^n], [q^\ell]$ representing here *formal* coefficient extraction in $\mathbb{C}[[z,q]]$. On the other hand, the *analytic* quantities $\mathcal{W}_\ell(z)$ are defined as

$$\mathcal{W}_\ell(z) = (-1)^{\ell+1}[q^{\ell+1}]\mathcal{Q}(z,q),$$

where the notation $[q^\ell]$ means now extraction of the coefficient of q^ℓ in the *asymptotic* expansion of $\mathcal{Q}(z,q)$ as a function of q with $q \to 0^+$. Thus, under the assumption (15), Theorem 1 only depends on the validity of the *interchange of coefficient operators*:

$$[q^\ell]\left([z^n]Q(z,q)\right) \overset{?}{=} [z^n]\left([q^\ell]\mathcal{Q}(z,q)\right).$$

Naturally, the divergent character of the underlying series renders this interchange non-obvious.

The basic ingredient is a lemma that grants conditionally such an interchange of limits and coefficient operators.

Lemma 3 (Interchange of limits and coefficients). (*i*) *Let $s_n(u)$ be polynomials with* nonnegative *coefficients and assume that the series*

$$S(z,u) = \sum_{n\geq 0} s_n(u) z^n$$

converges for $|z| < r$ (some $r > 0$) and $|u| < 1$. Assume that there exists a function $f(z) = \sum_n f_n z^n$ analytic in $|z| < r$ such that

$$\lim_{u \to 1^-} S(z,u) = f(z) \qquad \text{pointwise for any } z,\ |z| < r.$$

Then for all $n \geq 0$, the polynomials s_n converge:

$$s_n(1) \equiv \lim_{u\to 1^-}[z^n]S(z,u) = [z^n]\lim_{u\to 1^-} S(z,u) \equiv f_n.$$

(*ii*) *Additionally, assume that there exist functions $g_0(z) = f(z)$, $g_1(z), \ldots$ analytic in $|z| < r$ such that for $u \to 1^-$, one has*

$$S(z,u) \underset{u\to 1^-}{\sim} \sum_{k=0}^{\infty}(u-1)^k g_k(z).$$

Then, the derivatives of the polynomials s_n also converge, and, for all $k, n \geq 0$,

$$\left.\frac{d^k}{du^k} s_n(u)\right|_{u=1} = k![z^n]g_k(z).$$

Proof. (i) For any fixed u, write $S_u(z) = S(z,u)$ and consider the family of analytic functions $\{S_u(z)\}$ (in the variable z) indexed by u that ranges between 0 and 1 while tending to 1. Inside the disk $|z| < r$, the convergence $S_u(z) \to f(z)$ is dominated by $f(|z|)$: for positive z, this results plainly from the positivity of the $s_n(u)$, and for arbitrary z, from the triangular inequality. In particular, the convergence is dominated by a constant $M(r')$ in an arbitrarily chosen sub-disk $|z| \le r'$ with $0 < r' < r$. By a classical result of the theory of analytic functions, bounded pointwise convergence on compact sets implies uniform convergence. (See the discussion of *normal families* of functions in [16, Ch. 12] or [17, Ch. 15], especially pp. 246–247.) Thus, $S_u(z)$ converges to $f(z)$ uniformly in any sub-disk of $|z| < r$, so that, as $u \to 1^-$,

$$s_n(u) = \frac{1}{2\pi i}\int_{|z|=r'} S(z,u)\frac{dz}{z^{n+1}} \;\to\; f_n = \frac{1}{2\pi i}\int_{|z|=r'} f(z)\frac{dz}{z^{n+1}}.$$

This shows that $\lim_{u\to 1^-} s_n(u) = f_n$, and the form $s_n(1) = f_n$ of part (i) of the assertion follows by continuity of polynomials.

(ii) The proof follows by induction on k. Assume that $S(z,u)$ now satisfies the stronger conditions of (ii) and that the conclusion is met up to $k-1$. Set

$$T(z,u) = \frac{1}{(u-1)^k}\left(S(z,u) - \sum_{i=0}^{k-1} g_i(z)(u-1)^i\right).$$

Then, by assumption, $T(z,u)$ admits a shifted asymptotic expansion of the same type as $S(z,u)$, and in particular, it converges to the limit $g_k(z)$ as $u \to 1^-$. Moreover, one has

$$T(z,u) = \sum_{n\ge 0} t_n(u) z^n \quad \text{with} \quad t_n(u) = \frac{1}{(u-1)^k}\left(s_n(u) - \sum_{i=0}^{k-1} s_n^{(i)}(1)\frac{(u-1)^i}{i!}\right),$$

where the $t_n(u)$ are polynomials in u. Now, if the $s_n(u)$ have nonnegative coefficients, then so do the $t_n(u)$. Thus, part (i) of the statement applies to the function $T(z,u)$, giving

$$\frac{1}{k!}s_n^{(k)}(1) = \lim_{u\to 1} t_n(u) = [z^n]g_k(z),$$

so that the conclusion is satisfied for k. □

In summary, Lemma 3 asserts that, under suitable conditions,

$$[(u-1)^k]\left([z^n]S^{\text{ana}}(z,u)\right) = [z^n]\left([(u-1)^k]S^{\text{asy}}(z,u)\right).$$

There, the notations stress the fact that an object $S(z,u)$ is taken either as an analytic function S^{ana} at $(0,0)$ or as the corresponding asymptotic expansion S^{asy} as $u \to 1$. The coefficient notations $[(u-1)^k]$ are to be interpreted accordingly.

1.4. **Positivity of the Graphical Divergent Expansions.** Finally, as we show now, the function $Q(z,q)$ satisfies the conditions of Lemma 3. This corresponds to a supplementary positivity property, itself established by a specific *external* argument based on depth-first search traversal of graphs and inversions in trees [14].

Lemma 4. *Assume that, pointwise for each z with $|z| < e^{-1}$, and as $q \to 0^+$, the bivariate generating function of connected graphs $\mathcal{Q}(z,q)$ satisfies an asymptotic expansion,*

$$\mathcal{Q}(z,q) \underset{q\to 0^+}{\sim} \sum_{\ell\geq 0} \mathcal{W}_{\ell-1}(z)(-q)^{\ell}, \qquad |z| < e^{-1}, \tag{16}$$

for a sequence of functions $\mathcal{W}_\ell(z)$. Then, for each ℓ, the identity $\mathcal{W}_\ell(z) = W_\ell(z)$ holds, where $W_\ell(z)$ is defined algebraically by (11).

Proof. Let $U_{n,\ell}$ denote the number of unrooted labelled trees with n vertices and ℓ inversions[4] and let

$$U_n(u) := \sum_{\ell\geq 0} U_{n,\ell} u^{\ell} \qquad \text{and} \qquad U(z,u) := \sum_{n\geq 0} U_n(u)\frac{z^n}{n!},$$

be the corresponding generating functions. A tree with n vertices has at most $\binom{n-1}{2}$ inversions so that the $U_n(u)$ are polynomials; the polynomials also have positive coefficients given their combinatorial origin. Moreover $U(z,u)$ is analytic for $|z| < e^{-1}$ and $|u| < 1$ since the number of unrooted trees is *a priori* bounded from above by n^n. (A tree is specified by n daughter-to-mother links).

There is an elegant relation between inversions in unrooted trees and connected graphs discovered by Ira Gessel and Da Lun Wang [14] who proved combinatorially the formal power series relation $U(z, 1+q) = Q(z,-q)$. In essence, a connected graph may be considered as rooted at 1. From this root node, a depth first search traversal (with a suitable ordering on successor nodes) gives rise to a tree together with additional return edges that have to be to inversions. Conversely, each inversion in a tree may or may not be "activated" depending on the particular graph under consideration, and this fact is seen to be reflected by the relation $U(z,1+q) = Q(z,-q)$. (The negative argument $-q$ in Q is there since we adopted a negative variable to mark excess.) As a consequence, we have the fundamental relation

$$Q(z,q) = U(z,1-q). \tag{17}$$

By direct combinatorics, $U(z,u)$ has positive coefficients at $(0,0)$ and is bivariate analytic in $|z| < e^{-1}, |u| < 1$. Now an asymptotic expansion as $q \to 0^+$,

$$\mathcal{Q}(z,q) \underset{q\to 0^+}{\sim} \sum_{\ell\geq 0} \mathcal{W}_{\ell-1}(z)(-q)^{\ell},$$

can be recast via the relation (17) as an expansion of $U(z,u)$ as $u \to 1^-$,

$$U(z,u) \underset{u\to 1^-}{\sim} \sum_{\ell\geq 0} \mathcal{W}_{\ell-1}(z)(u-1)^{\ell}.$$

Now, Lemma 3 applies. Part (i) of Lemma 3 gives us already $W_0(z) = \mathcal{W}_0(z)$ through $[z^n]W_0(z) = [z^n]\mathcal{W}_0(z)$, this without any requirement other than (16). More generally, the equality $W_\ell(z) = \mathcal{W}_\ell(z)$ follows from Part (ii) of Lemma 3 via the identities $[z^n]W_\ell(z) = [z^n]\mathcal{W}_\ell(z)$ valid for all n. □

The discussion above allows us to identify Q and $\mathcal{Q}$. Accordingly, we shall use the notation Q in the rest of this article.

[4] An inversion is a pair of vertices (i,j) such that $1 < i < j$ and j is on the branch from i to 1.

2. Single Saddle-Point Analysis

We now proceed with the estimation of $Q(z,q)$ as $q \to 0^+$, starting from the integral representation (14) of Lemma 2. As will appear shortly, the "tree function" $T(z)$ of (2) is essential in our developments, and we shall accordingly adopt $t = T(z)$ as the main parameter (so that $z = te^{-t}$). In this section, the objective is to prove Wright's expansion (Theorem 1) by an analysis of the integral representation (14) when t is restricted to some fixed interval $(0, a)$ with $a < 1$. Precisely, the *single saddle-point* analysis of this section is summarized by an expansion (27) of the form

$$-\frac{1}{q}Q(z,q) + \frac{1}{q}W_{-1}(z) - W_0(z) \sim \sum_{k\geq 1} A_k(t)(-\alpha)^k \qquad (\alpha \to 0^+), \tag{18}$$

where we have set $t = T(z)$ and $\alpha = q/(1-t)^3$.

The analysis proceeds in four steps: first a modification of the integration contour in the representation of Lemma 2 in order to obtain a saddle-point representation; second, a standard change of variables in order to normalize the saddle-point integrand; third, formal termwise integration; fourth, an analysis of the remainder of the expansion in order to prove that the formal result is indeed an asymptotic expansion of the integral.

2.1. Saddle-point representation.

When $q \to 0^+$, the integrand in (14) oscillates more and more wildly, this because of the term $e^{ix\lambda}/q$ it contains. The tactics consist in disposing of the oscillation by shifting the integration contour so as to cross a saddle point. First, we set $x\lambda = w$, which transforms the integral into

$$Q(z,q) = -q \log \frac{1}{\lambda\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{w^2}{2\lambda^2} - z\frac{(1-q)^{-1/2}}{q}e^{iw}\right) dw.$$

The integrand rewrites as

$$\exp\left(-\frac{1}{q}\left(\frac{w^2}{2} + ze^{iw}\right)\right) \cdot \exp\left(\frac{w^2}{2}(q^{-1} - \lambda^{-2}) + ze^{iw}\frac{1-(1-q)^{-1/2}}{q}\right). \tag{19}$$

In this product, the first factor captures the dominant part of the integrand, while the second one acts as a small perturbation since it tends to a finite limit as $q \to 0^+$.

The saddle points ζ of the first factor are located at points ζ such that

$$\frac{d}{dw}\left(\frac{w^2}{2} + ze^{iw}\right)_{w=\zeta} \equiv \zeta + ize^{i\zeta} = 0. \tag{20}$$

We recall first some basic facts concerning the function $T(z)$ which is defined as the solution analytic at 0 of $T = ze^T$. On its radius of convergence $|z| = e^{-1}$, $T(z)$ has a unique singularity $z_0 = e^{-1}$ and $T(z_0) = 1$; moreover as z describes the real segment $(0, e^{-1})$, $T(z)$ increases from 0 to 1. We recognize in (20) the equation satisfied by $T(z)$, so that, as long as $|z| < e^{-1}$, we can take $t = T(z)$ and obtain a saddle point

$$\zeta = -it = -iT(z).$$

In subsequent computations, $t = T(z)$ is taken as the independent variable (rather than z itself), and is restricted when the need arises to be a real quantity in $(0,1)$. (Analytic continuation makes it possible to extend the domain of validity of end formulæ, if needed.)

The saddle-point method now suggests shifting the line of integration parallel to itself so that it crosses the point ζ. This does not change the value of the

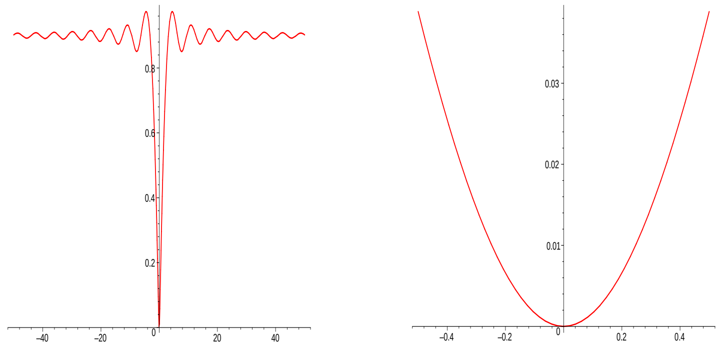


FIGURE 1. The steepest descent line $\Im(f(u)) = 0$ for $t = 0.5$: general aspect (left) and blow up near 0 (right).

integral by virtue of Cauchy's theorem and the fact that the integrand is small as $\Re(w) \to \pm\infty$. Thus, using $t = T(z)$ as a parameter, setting $w = u - it$ and replacing the integration contour on $(-\infty, +\infty)$ yields

$$\begin{aligned} Q(z,q) &= \left(t - \frac{t^2}{2}\right) + \left(1 - \frac{q}{\lambda^2}\right)\frac{t^2}{2} - q \log \frac{I}{\lambda\sqrt{2\pi}}, \\ I &:= \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{q}\left(\frac{u^2}{2} + t(e^{iu} - 1 - iu)\right)\right) h(u)\, du \end{aligned} \tag{21}$$

with

$$h(u) = \exp\left(\left(\frac{u^2}{2} - uit\right)(q^{-1} - \lambda^{-2}) + te^{iu}\frac{1 - (1-q)^{-1/2}}{q}\right).$$

The new integral form (21) "explains" the rôle of the tree function in the problem. In effect, it will turn out that the first term in (21) dominates as $q \to 0$, so that it provides the enumeration of unrooted trees, *i.e.*, Part (i) of Theorem 1.

2.2. **Change of variable.** First, we reduce the kernel of the saddle-point integral to standard *quadratic* form. The corresponding change of variable is defined by the equation

$$y^2 = f(u) \qquad \text{where} \quad f(u) := \frac{u^2}{2} + t(e^{iu} - 1 - iu). \tag{22}$$

We opt to perform the change of variable in such a way that y varies continuously on the real line from $-\infty$ to $+\infty$. Given the geometry of $f(u)$, this corresponds to taking the integral in (21) along the contour depicted in Figure 1—in fact a steepest descent line connecting $-\infty$ to $+\infty$. The value of the integral remains unaffected by virtue of analyticity and Cauchy's theorem.

The expression of the integral is changed into

$$I = \int_{-\infty}^{+\infty} e^{-y^2/q} H(y)\, dy, \qquad H(y) := h(u(y))\frac{du}{dy}. \tag{23}$$

2.3. **Term by term integration.** The next step is to expand H as a power series in y and integrate termwise. The validity of this process will be proved later. The net result of this formal manipulation is to effect a linear transformation $\mathfrak{L}$ on y-expansions: odd powers of y disappear while even powers are transformed by

$$\mathfrak{L}(y^{2k}) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-y^2/q} y^{2k}\, dy = \frac{1 \cdot 3 \cdots (2k-1)}{2^k} q^{k+1/2}.$$

In order to compute the series expansion of H, we examine the formulæ induced by the change of variables. The quantity $f(u)/u^2$ is an entire function of u, and

$$f(u) = (1-t)\frac{u^2}{2}\left(1 + \frac{2tu}{1-t}\frac{e^{iu} - 1 - iu + u^2/2}{u^3}\right).$$

The Taylor expansion of its square-root thus has the form

$$y(u) = (1-t)^{1/2}\frac{u}{\sqrt{2}}\sum_{k\geq 0} a_k(t)\left(\frac{u}{1-t}\right)^k,$$

with $a_0(t) = 1$ and $a_k(t)$ a polynomial of degree k in t. Thus $y(1-t)^{-3/2}$ has a Taylor expansion in powers of $u/(1-t)$ with coefficients that are polynomials in t. Then by reversion of formal power series, the expansion of $u(y)$ is of the form

$$u(y) = (1-t)\sum_{k\geq 1} b_k(t)\frac{y^k}{(1-t)^{3k/2}}, \tag{24}$$

with $b_1(t) = \sqrt{2}$ and coefficients $b_k(t)$ that are again polynomial in t, and the expansion has a positive radius of convergence. Composing expansions then yields

$$H(y) = \frac{\sqrt{2}\, h(0)}{(1-t)^{1/2}}\left(1 + \sum_{k\geq 1} c_k(t,q)\frac{y^k}{(1-t)^{3k/2}}\right), \tag{25}$$

where the coefficients $c_k(t,q)$ are computable polynomials in t which are analytic with respect to q for $|q| < 1$ as follows from the definition of h.

We have thus obtained the formal divergent expansion for Q:

$$\begin{aligned} Q(z,q) \underset{q\to 0^+}{\sim} &\left(t - \frac{t^2}{2}\right) + \left(1 - \frac{q}{\lambda^2}\right)\frac{t^2}{2} - \frac{q}{2}\log\frac{1}{1-t} - \frac{q}{2}\log\frac{q}{\lambda^2} \\ &- t(1-(1-q)^{-1/2}) - q\log\left(1 + \sum_{k\geq 1} c_{2k}(t,q)\frac{1 \cdot 3 \cdots (2k-1)}{2^k}\frac{q^k}{(1-t)^{3k}}\right), \end{aligned} \tag{26}$$

In particular, the scaling of the integral provides the enumeration of unicyclic graphs. Assuming the above formal expansion is asymptotic to Q (this is proved in §2.4 below), expansions at any finite order with respect to q are legitimate and yield finite order expansions for $q \to 0^+$ of the bivariate generating function of connected graphs counted by size and excess:

$$\begin{aligned} Q(z,q) \underset{q\to 0^+}{\sim} &\left(T(z) - \frac{T(z)^2}{2}\right) - \left(\frac{1}{2}\log\frac{1}{1-T(z)} - \frac{T(z)}{2} - \frac{T^2(z)}{4}\right)q \\ &+ \sum_{k\geq 2}\frac{A_{k-1}(T(z))}{(1-T(z))^{3k-3}}(-q)^k, \end{aligned} \tag{27}$$

where the A_k's are polynomials in $T(z)$.

This gives the results of Cayley and Rényi, as announced in (3), (4), as well as Wright's result as stated in (5).

This derivation provides a mechanical way to determine the $\mathcal{W}_k(z)$ by a simple process: (*i*) compute $u(y)$ by Eq. (22) and (24); (*ii*) determine the compound expansion (25); (*iii*) integrate termwise by (26) and conclude by expanding the logarithm like in (27). Barely a dozen instructions in a computer algebra system are needed to implement the algorithm. The computation yields in particular $A_1(t) = t^4(6-t)/(24(1-t)^3)$; a table of the first ten A_k (as a function of $\theta = 1-t$) is given in the appendix.

2.4. **Analysis.** To complete the proof of Theorem 1, we now legitimate term by term integration, thereby establishing that the right-hand side of (27) is an asymptotic expansion of Q for fixed real $t = T(z)$ in $(0,1)$. This is a variant of the classical Laplace method, where the c_k's depend on q, see also [25, p. 376].

Define the function $H_n(q,y)$ by (cf. (25))

$$H(y) = \frac{\sqrt{2}\,h(0)}{(1-t)^{1/2}}\left(1 + \sum_{k=1}^{n-1}\frac{c_k(t,q)y^k}{(1-t)^{3k/2}} + H_n(q,y)\right). \tag{28}$$

For fixed $n > 0$, integration termwise of the inner polynomial in y yields the initial part of the formal expansion of the integral I. To show that this process leads to an *asymptotic* expansion of I when $q \to 0$, it is sufficient to show that

$$\int_{-\infty}^{\infty} e^{-y^2/q} H_n(q,y)\,dy = O(q^{\frac{n+1}{2}}). \tag{29}$$

Since H and the coefficients c_k, $k = 1,\ldots,n-1$ are analytic with respect to q for $|q| < 1$, so is H_n. Moreover, H_n is analytic with respect to y in some neighbourhood of 0 since H is. The bound (29) is obtained using different tools inside and outside the disc of convergence of u defined by (24), whose radius we denote by R. Consider some $r_0 < R$. For real y with $|y| \geq r_0$, we have

$$|h(u(y))| \leq \exp(Cy^2)$$

where C does not depend on q, in view of the definition of h and the fact that $|y| \sim \sqrt{2}|u|$ for large y. From the change of variable (22), we also get for $|y| \to \infty$

$$\left|\frac{du}{dy}\right| = \left|\frac{2y}{f'(u)}\right| \sim \sqrt{2}|y|$$

so that by continuity there exists C' independent of q such that $|H(y)| \leq \exp(C'y^2)$ for $|y| \geq r_0$. Since H and H_n differ only by a polynomial, such a bound also holds for H_n. Therefore the portion of the integral (29) corresponding to $|y| \geq r_0$ is $O(\exp(-r_0^2/q))$ when $q \to 0$.

The bound for $|y| \leq r_0$ is obtained by first bounding the coefficients $d_k = c_k/(1-t)^{3k/2}$ uniformly with respect to q and then using a simple argument of majorizing series. Indeed, these coefficients are expressed by the Cauchy integral

$$d_k = \frac{1}{2i\pi}\oint \frac{F(q,y)}{y^{k+1}}\,dy,$$

where F is directly related to H and is an analytic function of q and y in $|q| < 1$ and $|y| < R$. A valid contour of integration is a circle of radius $r_0 + \delta < R$, on which $|F|$ is uniformly bounded with respect to q (for $|q| < 1/2$) by continuity. This shows that $d_k \leq C(r_0+\delta)^{-k}$ for some C that does not depend on q and k. From there follows that $|H_n(q,y)|$ is bounded by $C' r_0^{-n}|y|^n$ for some C' that does not depend on q whence the bound (29) which concludes the proof of Theorem 1.

3. Coalescing Saddle Points and the Airy Connection

In this section, we construct the generating function of the constants $A_k(1)$ that gives the dominant asymptotics of the number of connected graphs of a fixed excess (Theorem 2, Part (*i*) and Corollary 1). At the same time, we obtain an access to the successive "correction series" (Theorem 2, Part (*ii*)).

The *single saddle-point* analysis of the previous section is summarized by expansion (18) that we now recall

$$-\frac{1}{q}Q(z,q) + \frac{1}{q}W_{-1}(z) - W_0(z) \sim \sum_{k\geq 1} A_k(t)(-\alpha)^k \qquad (\alpha \to 0^+), \tag{30}$$

where we have set $t = T(z)$ and $\alpha = q/(1-t)^3$. Such an expansion holds for t in any closed subinterval of $[0,1)$, for instance $t \in [0, \frac{3}{4}]$, since it is nothing but a linear rescaling in the asymptotic variable, as long as t avoids 1. However, the expression becomes meaningless, should t approach 1. Accordingly, the proof of (30) given in Section 2 gives access to successive *lower order terms* of the polynomials $A_k(t)$ near $t = 0$.

Here, we develop a more sophisticated analysis based on a method of *coalescent saddle points* whose principles originate with Chester, Friedman, and Ursell [6] and which is exposed in classical treatises like [4, 25, 33]. In particular, we follow closely the treatment offered by Olver in [25, p. 352–356]. Proceeding along these lines, we establish below the existence of an expansion

$$-\frac{1}{q}Q(z,q) + \frac{1}{q}W_{-1}(z) - W_0(z) \sim \sum_{k\geq 1} B_k(t,\alpha) \qquad (\alpha \to 0^+), \tag{31}$$

that is valid for t in a closed subinterval of $(0,1]$, for instance $t \in [\frac{1}{4}, 1]$. It will appear that, for t fixed, as α tends to 0, the $B_k(t,\alpha)$ form a proper asymptotic scale with B_k being proportional to α^k as $\alpha \to 0+$. It will also appear that each B_k admits an expansion in α, as α tends to 0,

$$B_k(t,\alpha) \sim \sum_{i\geq k} b_{k,i}(t)\alpha^i,$$

where additionally the $b_{k,\ell}(t)$ are polynomials in t. Once this is granted, the expansion (31) can be "finitely" reorganized into powers of α, and, given that (30) and (31) have a common domain of validity (say $t \in [\frac{1}{4}, \frac{3}{4}]$), uniqueness of asymptotic expansions in the scale of powers of α shows that

$$(-1)^k A_k(t) = b_{1,k}(t) + b_{2,k}(t) + \cdots + b_{k,k}(t).$$

Naturally, such an equality that holds numerically for values of t in $[\frac{1}{4}, \frac{3}{4}]$ lifts to an identity between polynomials. Setting $\theta = 1 - t$, this allows us to catch the generating functions of derivatives of the $A_k(t)$,

$$\sum_{k\geq 1} A_k^{(j)}(1)\,(-\alpha)^k \sim (-1)^j j! \sum_{\ell\geq 1} [\theta^j] B_\ell(t,\alpha),$$

where we obtain the coefficients of θ^j through expansions at $\theta \to 0$, and see that only a finite number of $B_\ell(t,\alpha)$ contributes to θ^j for a given j. The generating functions of derivatives are thus seen to be accessible to analysis starting from terms of *lower order* in θ from (31).

Throughout the rest of the section, we use the notations

$$\theta = 1 - t \qquad \text{and} \qquad \alpha = \frac{q}{(1-t)^3},$$

and think of α as a quantity that tends to 0 while θ is a parameter that ranges over some arbitrary interval $(0, b]$, with 0 being allowed in the limit.

The treatment offered here follows closely in the steps of the single saddle-point analysis of Section 2, with a different change of variables: first an analysis of the location of dominant saddle points; second, a *cubic* change of variables in order to normalize the saddle-point integrand; third, formal termwise integration; fourth, an analysis of the remainder of the expansion in order to prove that the formal result is indeed an asymptotic expansion of the integral.

3.1. **Saddle points.** The starting point is the integral representation (21):

$$(32) \qquad I := \int_{-\infty}^{+\infty} e^{-f(u)/q} h(u)\, du, \quad \text{with} \quad f(u) = \frac{u^2}{2} + (1-\theta)\left(e^{iu} - 1 - iu\right).$$

The quantity h preserves its former meaning, but now with $q = \alpha\theta^3$.

The main result of the previous section (Eq. (27)) is of a form (30) that ceases to be valid when θ approaches 0 (*i.e.*, $t \to 1$). One reason is that $f(u)$ becomes locally cubic at $\theta = 0$ instead of being quadratic when $\theta \neq 0$. Solving $f'(u) = 0$ for $u \neq 0$ in the neighbourhood of the origin reveals a "shadow" saddle point ρ that is purely imaginary and satisfies the expansion

$$(33) \qquad \rho = -2i\theta\left(1 + \frac{1}{3}\theta + \frac{2}{9}\theta^2 + \frac{22}{135}\theta^3 + \cdots\right).$$

Non-uniformity arises precisely from the coalescence of the two nearby saddle points at 0 and ρ, as $t \to 1$. By construction, $f(0) = 0$, while the value of f at the other saddle point is

$$f(\rho) = -\frac{2}{3}\theta^3 - \frac{2}{3}\theta^4 - \frac{28}{45}\theta^5 - \cdots.$$

We make use of several expansions related to ρ, as well as the position of saddle points other than 0 and ρ. A convenient expression for such quantities is provided through the use of the indexed Lambert $\mathbf{W}$ function [7]. This function is solution of $ye^y = z$. It is a multivalued function and the index is used to distinguish between different branches. Thus the first saddle point $-it$ introduced in Section 2 is $-it = i\mathbf{W}_0(-z)$, while

$$\rho = i(\mathbf{W}_{-1}(-z) - \mathbf{W}_0(-z)) = i\left(t + \mathbf{W}_{-1}(-te^{-t})\right) = i\left(1 - \theta + \mathbf{W}_{-1}((\theta-1)e^{\theta-1})\right).$$

Of course, the $\mathbf{W}_j$ bear no relation to Wright's generating functions W_k or $\mathcal{W}_k$.

As illustrated by Figure 2, the next closest saddle points are two symmetrical points σ and $-\bar{\sigma}$, that are given by

$$(34) \qquad \sigma = i(1 - \theta + \mathbf{W}_{-2}((\theta-1)e^{\theta-1})).$$

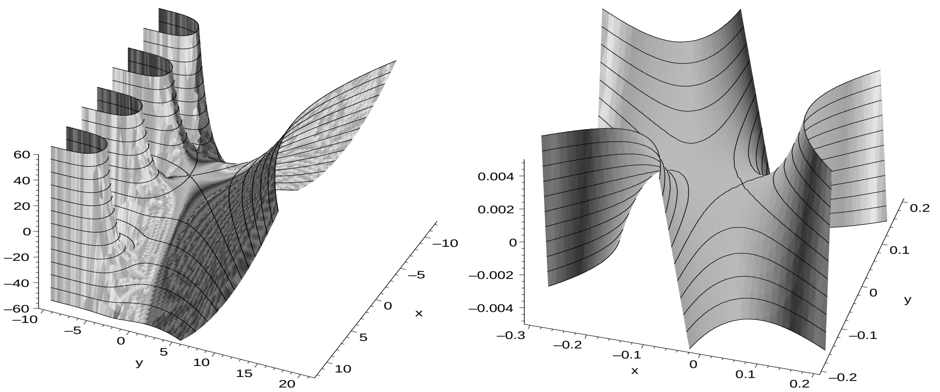


FIGURE 2. The landscapes of $-\Re f(u)$ and $-\Re P(v)$.

3.2. **Change of variable.** In order to estimate asymptotically the integral in (32), the classical method of Chester, Friedmann and Ursell (see in particular [25, p. 352–356]) is used. Consequently, we introduce the *cubic* change of variable

$$f(u) = P(v) \qquad \text{where} \quad P(v) = \frac{f(\rho)}{\theta^3}(2v^3 + 3\theta v^2). \tag{35}$$

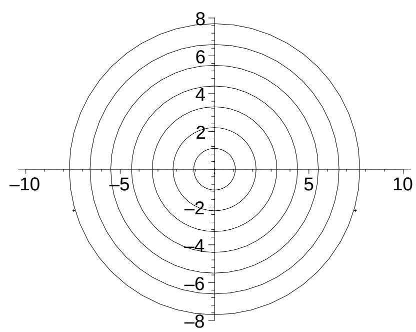
The polynomial P is such that P' has two roots at 0 and $-\theta$, $P(0) = 0$ and $P(-\theta) = f(\rho)$. Thus P and f behave similarly in the neighbourhood of their two central saddle points, and one expects the change of variable to be conformal in this neighbourhood. Indeed, as illustrated by Figure 2, it is only when approaching the next saddle points σ and $-\bar{\sigma}$ of f that the two landscapes start to diverge qualitatively. Numerical experiments indicate that the change of variables is one-to-one for any $0 < \theta < 1$ and $|u| < |\sigma(0)| = |1 + \mathbf{W}_{-2}(-e^{-1})| \approx 7.748360311$. However, in our proofs it will be sufficient to make use of the following.

Lemma 5. [6, Th. 1] *There exists $\theta_0 > 0$ and $r_u > 0$ such that the change of variable (35) is one-to-one for any (θ, u) such that $|\theta| < \theta_0$, $|u| < r_u$.*

Since f does not possess any saddle point outside 0 on the real axis, the change of variable is also one-to-one on the whole domain of integration. The new contour of integration is obtained by following consistently the proper branch of the cubic (35): for real u with large absolute value, f is positive; since $f(\rho) < 0$, this forces $\lim \arg(v) = \pm\pi/3$; for small u, $f(u) \sim \theta u^2/2$ so that the contour is vertical in the neighbourhood of $v = 0$; finally, $v = -\theta$ corresponding to $u = \rho$ fixes the orientation on the contour. The integral (32) thus admits the exact expression:

$$I = -\int_{e^{-i\pi/3}\infty}^{e^{i\pi/3}\infty} e^{-P(v)/q} G(v)\, dv, \qquad G(v) = h(u(v))\frac{du}{dv}. \tag{36}$$

Figure 3 displays the images of circles and of the real axis by the change of variable. Small points on the left correspond to the next saddle points σ and $-\overline{\sigma}$, that are mapped to cusps on the right in the v-plane.



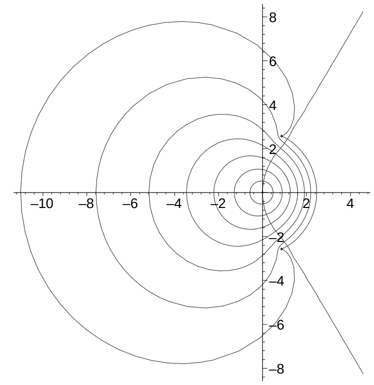


FIGURE 3. Circles and real axis in the u-plane (left) and their images in the v-plane (right) for $\theta = 1/10$.

3.3. **Term by term integration.** As in the single saddle-point analysis, the next step consists in expanding G as a power series

$$G(v,\alpha,\theta) = \sum_{k\geq 0} g_k(\alpha,\theta)v^k, \tag{37}$$

and integrating termwise. The validity of the process is proved later. This formal manipulation reduces to a linear transformation $\mathfrak{M}$ on v-expansions defined by

$$\mathfrak{M}[\phi] := \int_{e^{-i\pi/3}\infty}^{e^{i\pi/3}\infty} e^{-P(v)/q}\phi(v)\,dv.$$

The transformation involves the basic integrals

$$R_k(\xi) := \int_{e^{-i\pi/3}\infty}^{e^{i\pi/3}\infty} e^{\xi^{-1}(2v^3+3v^2)}v^k\,dv. \tag{38}$$

The net result of this formal step is an expansion of the fundamental integral I, namely

$$I \sim -\sum_{k\geq 0} g_k(\alpha,\theta)\theta^{k+1}R_k\left(\frac{-\theta^3}{f(\rho)}\alpha\right). \tag{39}$$

We show in the next section that $R_k(\xi)$ behaves roughly like $(\xi^{1/2}/\sqrt{3})^k$ when $\xi \to 0$. We know that $Q(z,q)$ is determined principally by $\log I$. At this stage, it suffices to compose expansions in a routine way in order to obtain the final expansion of $Q(z,\alpha\theta^3)$ as $\alpha \to 0$.

The rest of this section completes the proof of Theorem 2 by giving:

- — precise information on the basic quantities R_k, including the fact that these form an asymptotic scale;
- — the actual computation of the series expansion (37) and of the resulting expansion of $Q(z,q)$;
- — an argument ensuring the validity of term by term integration, which conditions the final result.

3.3.1. *The hypergeometric scale.* The asymptotic expansion (39) involves the basic quantities $R_k(\xi)$ of which the character as $\xi \to 0$ is needed. The expressions to be obtained involve the generalized hypergeometric series that are classically defined [8] by

$$ {}_pF_q\left({a_1, \ldots, a_p \atop b_1, \ldots, b_q} \,\middle|\, x \right) := \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{x^n}{n!}, $$

where $(a)_n = a(a+1)\cdots(a+n-1)$. For $p > q+1$, such series have radius of convergence zero. However, they can still be used to express asymptotic expansions, as we see now.

Lemma 6. *For real $\xi \to 0^+$, R_k admits the following (divergent) power series expansion:*

$$ R_k(\xi) \underset{\xi\to 0}{\sim} \begin{cases} i(-1)^{\frac{k}{2}} \left(\frac{\xi}{3}\right)^{\frac{k+1}{2}} \Gamma\left(\frac{k+1}{2}\right) {}_3F_1\left({\frac{k+1}{6}, \frac{k+3}{6}, \frac{k+5}{6} \atop 1/2} \,\middle|\, -\xi \right), & k \text{ even}, \\ i(-1)^{\frac{k-1}{2}} \frac{2}{3} \left(\frac{\xi}{3}\right)^{\frac{k+2}{2}} \Gamma\left(\frac{k+4}{2}\right) {}_3F_1\left({\frac{k+4}{6}, \frac{k+6}{6}, \frac{k+8}{6} \atop 3/2} \,\middle|\, -\xi \right), & k \text{ odd}. \end{cases} $$

Moreover, the R_k are all reducible to linear combinations of

$$ A(\xi) = {}_2F_0\left({\frac{1}{6}, \frac{5}{6} \atop -} \,\middle|\, -\xi \right) \qquad \text{and} \qquad B(\xi) = {}_2F_0\left({\frac{5}{6}, \frac{7}{6} \atop -} \,\middle|\, -\xi \right). $$

These hypergeometric series are often encountered in uniform asymptotic expansions. In particular, A occurs in the asymptotic expansion of the classical Airy function [1, §10.4.59], while both A and B appear in the expansion of its derivative.

Proof. For real positive ξ, deforming the contour to the imaginary axis shows that another expression for R_k is

$$ R_k(\xi) = 2i \int_0^{+\infty} e^{-\frac{3v^2}{\xi}} v^k \cos\left(2\frac{v^3}{\xi} - k\frac{\pi}{2} \right) dv. $$

The change of variables $3v^2/\xi = w$ leads to

$$ 2i \left(\frac{\xi}{3}\right)^{\frac{k+1}{2}} \int_0^{+\infty} e^{-w} w^{k-1/2} \cos\left(\frac{2\xi^{1/2} w^{3/2}}{3^{3/2}} - k\frac{\pi}{2} \right) dw. $$

The results follows from expanding the cosine and integrating termwise, which is justified by Watson's lemma (see *e.g.* [16, Vol. II, p. 389]).

The reduction to $A(\xi)$ and $B(\xi)$ follows from contiguity relations. If E_k and O_k denote the hypergeometric series involved in the even and odd case respectively, then it can be seen by series manipulations (or by integration by parts) that

$$ \begin{aligned} \xi(k+3)(k+5)E_{k+6} + 6(k+3)E_{k+4} - 6(2k+7)E_{k+2} + 6(k+4)E_k &= 0, \\ \xi(k+6)(k+8)O_{k+6} + 6(k+6)O_{k+4} - 6(2k+7)O_{k+2} + 6(k+1)O_k &= 0. \end{aligned} $$

It is then sufficient to reduce the first three E_k's and O_k's to the desired form. The following can be proved by series expansion:

$$ \begin{array}{lll} E_0 = A(\xi), & E_2 = B(\xi), & E_4 = \frac{2}{\xi}A(\xi) - \frac{2}{3\xi}(3+\xi)B(\xi) \\ O_1 = B(\xi), & O_3 = \frac{6}{5\xi}(A(\xi) - B(\xi)), & O_5 = -\frac{36}{35\xi^2}A(\xi) + \frac{6}{35\xi^2}(6+5\xi)B(\xi). \end{array} $$

□

3.3.2. *Expansions.* Locally, a series expansion for the change of variables is obtained as in the previous section by taking square roots and inverting power series. This yields

$$
\begin{aligned}
u &= 2i\left(\sqrt{-\frac{3f(\rho)}{2\theta^3}}v + \left(\frac{2\sqrt{-\frac{3f(\rho)}{2\theta^3}}}{\theta} - 3\frac{f(\rho)}{\theta^3} + 3\frac{f(\rho)}{\theta^4}\right)\frac{v^2}{6} + \dots\right) \\
&= iv\left(2 + \theta + \frac{41}{60}\theta^2 + \cdots\right) + iv^2\left(\frac{1}{3} + \frac{49}{180}\theta + \cdots\right) + \cdots .
\end{aligned} \tag{40}
$$

Substituting this expansion in G yields (37) with first terms:

$$
\begin{aligned}
G(v) &= 2ie^{h(0)}(1 + (\frac{2}{3\theta}\sqrt{-\frac{3f(\rho)}{2\theta^3}} - 6\frac{f(\rho)}{2\theta^3}(1-\theta)(\frac{1}{q\sqrt{1-q}} - \frac{1}{\ln\frac{1}{1-q}} - \frac{1}{3}))v + O(v^2)) \\
&= 2ie^{-1/2}\left(1 + \theta + \frac{43}{60}\theta^2 + (\frac{277}{540} - \frac{3}{8}\alpha)\theta^3 + \cdots\right. \\
&\qquad \left. + v\left(\frac{7}{3} + \frac{259}{180}\theta + \frac{1673}{3240}\theta^2 + \left(\frac{147509}{907200} + \frac{1}{24}\alpha\right)\theta^3 + \cdots\right) + O(v^2)\right)
\end{aligned}
$$

The expansion of Q itself is obtained in the standard way from

$$
\begin{aligned}
Q(z,q) \sim \left(t - \frac{t^2}{2}\right) &+ \left(1 - \frac{\alpha\theta^3}{\lambda(\alpha\theta^3)^2}\right)\frac{t^2}{2} \\
&- \alpha\theta^3\log\left(\frac{-1}{\lambda(\alpha\theta^3)\sqrt{2\pi}}\sum_{k\geq 0} g_k(\alpha,\theta)\theta^{k+1}R_k\left(-\alpha\theta^3/f(\rho)\right)\right),
\end{aligned} \tag{41}
$$

by replacing λ, $f(\rho)$, the g_k and R_k by their expansion and reordering the terms in increasing powers of θ. This yields

$$
\begin{aligned}
\frac{1}{\sqrt{2\pi}\lambda}I = e^{-1/2}\theta^{-1/2}&\left({}_2F_0\left(\left.{\frac{1}{6},\frac{5}{6} \atop -}\right| -\frac{3\alpha}{2}\right)\right. \\
&\left. + \frac{\theta}{12}\left(8\,{}_2F_0\left(\left.{\frac{1}{6},\frac{5}{6} \atop -}\right| -\frac{3\alpha}{2}\right) + (7\alpha-2)\,{}_2F_0\left(\left.{\frac{5}{6},\frac{7}{6} \atop -}\right| -\frac{3\alpha}{2}\right)\right) + \cdots\right).
\end{aligned}
$$

Or, put otherwise:

$$
\begin{aligned}
Q(z,q) - (t - t^2/2) + \left(\frac{1}{2}\log\frac{1}{1-t} - \frac{t}{2} - \frac{t^2}{4}\right)q &= -\alpha\theta^3\ln\left({}_2F_0\left(\left.{\frac{5}{6},\frac{1}{6} \atop -}\right| -\frac{3\alpha}{2}\right)\right) \\
&- \frac{\theta^4}{12}(2\alpha + \alpha(7\alpha-2)S) \\
&+ \frac{\theta^5}{360}\left(\frac{5}{4}\alpha(7\alpha-2)^2S^2 + (245\alpha^2 - 94\alpha + 20)S + 114\alpha - 20\right) \\
&- \frac{\theta^6}{20160}\left(\frac{35}{9}\alpha(7\alpha-2)^3S^3 + \frac{14}{3}(7\alpha-2)(245\alpha^2 - 94\alpha + 20)S^2\right. \\
&\quad \left. + \frac{56}{9\alpha}(2019\alpha^3 - 840\alpha^2 + 144\alpha - 40)S - \frac{56}{9\alpha}(733\alpha^3 - 854\alpha^2 + 164\alpha - 40)\right) + \cdots,
\end{aligned}
$$

where

$$S = \frac{{}_2F_0\left(\begin{matrix}\frac{5}{6},\frac{7}{6}\\ -\end{matrix}\middle| -\frac{3\alpha}{2}\right)}{{}_2F_0\left(\begin{matrix}\frac{1}{6},\frac{5}{6}\\ -\end{matrix}\middle| -\frac{3\alpha}{2}\right)} = -\frac{2}{\alpha}\left(1+(2\alpha)^{1/3}\frac{\mathrm{Ai}'((2\alpha)^{-2/3})}{\mathrm{Ai}((2\alpha)^{-2/3})}\right) = 1+\frac{95}{288}\alpha+O(\alpha^2).$$

For $j \geq 0$, the coefficient of θ^{3+j} in this expansion is precisely the generating function of the $A_k^{(j)}(1)$. The first 5 generating functions are given in Appendix II. As a final check, expanding the coefficient of θ^4 in the above expansion of Q with respect to α, one gets

$$-\frac{19}{24}\alpha^2+\frac{65}{48}\alpha^3-\frac{1945}{384}\alpha^4+\frac{21295}{768}\alpha^5-\frac{603965}{3072}\alpha^6+\frac{10454075}{6144}\alpha^7+O\left(\alpha^8\right),$$

from which it is easy to recognize the coefficients of θ^2 in the polynomials A_k of Appendix I, that were obtained by the single saddle-point expansion.

3.4. **Analysis.** So far, we have proceeded formally without paying attention to convergence. We now show that the series (39) is asymptotic to the integral I. The proof is similar to that of §2.4, but is technically more demanding because of uniformity issues.

Recall that the cubic change of variables (35) is one-to-one for $|\theta| < \theta_0$ and either $|u| < r_u$ or u is real. Moreover, we shall make use of the following, where $r_v = \max_{|u|=r_u} |v(u)|$ and $\alpha_0 = \theta_0^{-3}$.

Lemma 7. *For $|\theta| < \theta_0$, $|\alpha| < \alpha_0$, $|v| < r_v$, the function $G(v,\alpha,\theta)$ is an analytic function of its arguments.*

Proof. The change of variables is analytic and does not involve α. By definition, $G(v,\alpha,\theta) = h(u(v))du/dv$. Its expression shows that h is analytic in u, θ and $q = \alpha\theta^3$ provided $|q| < 1$, which concludes the proof. □

We now define $G_n(\alpha,\theta,v)$ by (cf. (37) and the similar (28))

$$G(v) = \sum_{k=0}^{n-1} g_k(\alpha,\theta)v^k + v^n G_n(\alpha,\theta,v). \tag{42}$$

The proof that (39) is asymptotic to I is concluded by the following.

Lemma 8. *Let θ be $[0,\theta_0)$ and $\alpha > 0$, then*

$$I_n(\alpha,\theta) = \int_{e^{-i\pi/3}\infty}^{e^{i\pi/3}\infty} e^{-P(v)/q} v^n G_n(\alpha,\theta,v)\,dv = \theta^n O(\alpha^{n/2}). \tag{43}$$

As a consequence of the previous lemma, the coefficients $g_k(\alpha,\theta)$ in (37) are analytic for $|\theta| < \theta_0$ and $|\alpha| < \alpha_0$. The uniform bound on I_n then legitimates expanding the coefficients in (39) with respect to α and reorganizing the truncated series. Moreover, analyticity with respect to θ legitimates expanding with respect to θ, leading to an expansion whose coefficients are asymptotic expansions in α that give the generating series of the numbers $A_k^{(j)}(1)$ of Theorem 2.

Proof. As in the single saddle-point analysis, the proof of the bound on I_n is obtained by bounding $|G_n(\alpha,\theta,v)|$ in two different regions. For small v we compute a bound on the coefficients g_k and then use a majorizing series argument. For larger v we consider the behaviour of $G(\alpha,\theta,v)$ when $|v|$ is large but remains on the contour.

The coefficients g_k are given by Cauchy's formula

$$g_k(\alpha,\theta) = \frac{1}{2i\pi}\oint \frac{G(\alpha,\theta,v)}{v^{k+1}}\,dv,$$

where the contour is for instance a circle centered at the origin with radius $r \le r_v$. For $|\alpha| < \alpha_0$ and $\theta < \theta_0$, G being analytic is uniformly bounded and thus there is a constant M such that $|g_k(\alpha,\theta)| \le M/r^k$. From there it follows that

$$|G_n(\alpha,\theta,v)| \le M\frac{r^{-n+1}}{r-|v|},$$

as long as $|v| < r$.

In order to bound (43), we deform the contour into three pieces: a vertical segment from $-ir_v/2$ to $ir_v/2$; arcs of the circle $|v| = r_v/2$ from the extremities of this segment to the original contour of (43); the rest of this contour to infinity. The integral is then bounded on these three pieces separately.

On the vertical segment, the integral is bounded by

$$\left|\int_{-r_v/2\theta}^{r_v/2\theta} e^{3\frac{f(\rho)}{\alpha\theta^3}w^2}\theta^{n+1}w^n M r_v^{-n}\,dw\right| \le M r_v^{-n}\theta^{n+1}\left(\frac{-\alpha\theta^3}{3f(\rho)}\right)^{\frac{n+1}{2}} n!.$$

On the arcs of circle, a direct computation leads to the bound

$$\left|\int_{\pi/3}^{\pi/2} \exp\left(-\frac{f(\rho)}{4q\theta^3}(r_v^3\cos(3\phi)+3\theta r_v^2\cos(2\phi))\right)\left(\frac{r_v}{2}\right)^{n+1} 2Mr_v^{-n}\,d\phi\right| \le 2^{-n}\exp\left(\frac{Kf(\rho)}{\alpha\theta^5}\right),$$

where K is a positive constant. (The first cosine is negative and the second one is upper bounded by $-1/2$.)

For $|v| \ge r_v/2$, using the change of variables (35) and a reasoning similar to that of Section 2.4, we get that $|G_n(\alpha,\theta,v)| < \exp(C|P(v)|)$, where C is a positive constant. Injecting this bound into

$$\int_{r_0}^{+\infty} e^{-f(u)/q}G_n(v(u))\frac{dv}{du}\,du,$$

where r_0 is such that $|v(r_0)| = r_v/2$ leads to a bound $\exp(-C'r_0^2/\alpha\theta^3)$ for the remaining part of the integral. Since r_0 is bounded from below uniformly with respect to θ, this concludes the proof of the lemma. □

Acknowledgement. We are thankful to Gaël Rémond for suggesting that a positivity property might help in Section 1. This work was supported in part by the ALCOM-FT project (contract IST-1999-14186) of the European Union.

References

[1] Abramowitz, M., and Stegun, I. A., Eds. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Dover Publications Inc., New York, 1992. Reprint of the 1972 edition.

[2] Banderier, C., Flajolet, P., Schaeffer, G., and Soria, M. Random maps, coalescing saddles, singularity analysis, and Airy phenomena. *Random Structures & Algorithms 19*, 3/4 (2001), 194–246.

[3] Bender, E. A., Canfield, E. R., and McKay, B. D. The asymptotic number of labeled graphs with given number of vertices and edges. *Random Structures & Algorithms 1*, 2 (1990), 127–169.

[4] Bleistein, N., and Handelsman, R. A. *Asymptotic expansions of integrals*, second ed. Dover Publications Inc., New York, 1986.

[5] BOLLOBÁS, B. *Random Graphs.* Academic Press, 1985.

[6] CHESTER, C., FRIEDMAN, B., AND URSELL, F. An extension of the method of steepest descents. *Proceedings of the Cambridge Philosophical Society 53* (1957), 599–611.

[7] CORLESS, R. M., GONNET, G. H., HARE, D. E. G., JEFFREY, D. J., AND KNUTH, D. E. On the Lambert W function. *Advances in Computational Mathematics 5*, 4 (1996), 329–359.

[8] ERDÉLYI, A., MAGNUS, W., OBERHETTINGER, F., AND TRICOMI, F. G. *Higher transcendental functions. Vol. I.* Robert E. Krieger Publishing Co. Inc., Melbourne, Fla., 1981. Reprint of the 1953 original.

[9] FLAJOLET, P., GRABNER, P. J., KIRSCHENHOFER, P., AND PRODINGER, H. On Ramanujan's Q-function. *Journal of Computational and Applied Mathematics 58*, 1 (1995), 103–116.

[10] FLAJOLET, P., KNUTH, D. E., AND PITTEL, B. The first cycles in an evolving graph. *Discrete Mathematics 75* (1989), 167–215.

[11] FLAJOLET, P., AND LOUCHARD, G. Analytic variations on the Airy distribution. *Algorithmica 31*, 3 (2001), 361–377.

[12] FLAJOLET, P., AND NOY, M. Analytic combinatorics of chord diagrams. In *Formal Power Series and Algebraic Combinatorics* (2000), D. Krob, A. A. Mikhalev, and A. V. Mikhalev, Eds., pp. 191–201. Proceedings of the 12th International Conference, FPSAC'2000; June 2000, Moscow.

[13] FLAJOLET, P., POBLETE, P., AND VIOLA, A. On the analysis of linear probing hashing. *Algorithmica 22*, 4 (Dec. 1998), 490–515.

[14] GESSEL, I., AND WANG, D. L. Depth-first search as a combinatorial correspondence. *Journal of Combinatorial Theory,* Series A *26*, 3 (1979), 308–313.

[15] HARARY, F., AND PALMER, E. M. *Graphical Enumeration.* Academic Press, 1973.

[16] HENRICI, P. *Applied and Computational Complex Analysis.* John Wiley, New York, 1977. 3 volumes.

[17] HILLE, E. *Analytic function theory.* Ginn and Company, Boston, 1959. 2 volumes.

[18] JANSON, S., KNUTH, D. E., ŁUCZAK, T., AND PITTEL, B. The birth of the giant component. *Random Structures & Algorithms 4*, 3 (1993), 233–358.

[19] KNUTH, D. E. *The Art of Computer Programming*, 3rd ed., vol. 1: Fundamental Algorithms. Addison-Wesley, 1997.

[20] KNUTH, D. E. Linear probing and graphs. *Algorithmica 22*, 4 (Dec. 1998), 561–568.

[21] KNUTH, D. E., AND PITTEL, B. A recurrence related to trees. *Proceedings of the American Mathematical Society 105*, 2 (Feb. 1989), 335–349.

[22] LOUCHARD, G. The Brownian excursion: a numerical analysis. *Computers and Mathematics with Applications 10*, 6 (1984), 413–417.

[23] MALLOWS, C. L., AND RIORDAN, J. The inversion enumerator for labeled trees. *Bulletin of the American Mathematical Society 1968* (74), 92–94.

[24] MONASSON, R. Statistical physics of the random graph model. In *Algorithms Seminar 97-98* (1998), B. Salvy, Ed., vol. 3504 of *Inria Research Report*, pp. 57–62. Summary by Philippe Flajolet.

[25] OLVER, F. W. J. *Asymptotics and Special Functions.* Academic Press, 1974.

[26] RIDDELL, JR., R. J., AND UHLENBECK, G. E. On the theory of the virial development of the equation of state of mono-atomic gases. *Journal of Chemical Physics 21* (1953), 2056–2064.

[27] SPENCER, J. Enumerating graphs and Brownian motion. *Communications on Pure and Applied Mathematics 50* (1997), 293–296.

[28] TAKÁCS, L. A Bernoulli excursion and its various applications. *Advances in Applied Probability 23* (1991), 557–585.

[29] TAKÁCS, L. Conditional limit theorems for branching processes. *Journal of Applied Mathematics and Stochastic Analysis 4*, 4 (1991), 263–292.

[30] TAKÁCS, L. On a probability problem connected with railway traffic. *Journal of Applied Mathematics and Stochastic Analysis 4*, 1 (1991), 1–27.

[31] TEMPERLEY, H. N. V. On the enumeration of the Mayer cluster integrals. *Proc. Phys. Soc. Sect. B. 72* (1959), 1141–1144.

[32] WILF, H. S. *Generatingfunctionology.* Academic Press, 1990.

[33] WONG, R. *Asymptotic approximations of integrals.* Academic Press Inc., Boston, MA, 1989.

[34] WRIGHT, E. M. The number of connected sparsely edged graphs. *Journal of Graph Theory 1* (1977), 317–330.

[35] Wright, E. M. The number of connected sparsely edged graphs. II. Smooth graphs. *Journal of Graph Theory 2* (1978), 299–305.

[36] Wright, E. M. The number of connected sparsely edged graphs. III. Asymptotic results. *Journal of Graph Theory 4* (1980), 393–407.

Appendix I: Polynomials A_k for $k = 1, \ldots, 10$

The numerator polynomials of the generating function (5) of connected graphs with excess k. In this list, $\theta = 1 - t$.

$$A_1 = \frac{1}{24}(5 - 19\theta + 26\theta^2 - 14\theta^3 + \theta^4 + \theta^5)$$

$$A_2 = \frac{1}{48}(15 - 65\theta + 108\theta^2 - 87\theta^3 + 42\theta^4 - 23\theta^5 + 12\theta^6 - \theta^7 - \theta^8)$$

$$\begin{aligned}A_3 = \frac{1}{5760}&(5525 - 29175\theta + 63530\theta^2 - 74560\theta^3 + 53574\theta^4 - 27378\theta^5 + 11504\theta^6 - 4020\theta^7\\ &+ 1725\theta^8 - 879\theta^9 + 78\theta^{10} + 76\theta^{11})\end{aligned}$$

$$\begin{aligned}A_4 = \frac{1}{11520}&(50850 - 319425\theta + 861685\theta^2 - 1315525\theta^3 + 1277185\theta^4 - 860612\theta^5 + 441526\theta^6 - 185786\theta^7\\ &+ 66964\theta^8 - 21977\theta^9 + 6577\theta^{10} - 2481\theta^{11} + 1241\theta^{12} - 114\theta^{13} - 108\theta^{14})\end{aligned}$$

$$\begin{aligned}A_5 = \frac{1}{2903040}&(78269625 - 570746925\theta + 1833592950\theta^2 - 3431917090\theta^3 + 4195202095\theta^4 - 3596232423\theta^5\\ &+ 2302080676\theta^6 - 1170871408\theta^7 + 497513283\theta^8 - 183003459\theta^9 + 60117702\theta^{10} - 18042570\theta^{11} + 5147401\theta^{12}\\ &- 1399153\theta^{13} + 485184\theta^{14} - 239044\theta^{15} + 22444\theta^{16} + 20712\theta^{17})\end{aligned}$$

$$\begin{aligned}A_6 = \frac{1}{5806080}&(1189944000 - 9879100875\theta + 36778793625\theta^2 - 81347450975\theta^3 + 119775583445\theta^4\\ &- 125345458455\theta^5 + 98004064025\theta^6 - 59998295119\theta^7 + 30081596601\theta^8 - 12822495201\theta^9 + 4781320559\theta^{10}\\ &- 1596163521\theta^{11} + 486019607\theta^{12} - 137459453\theta^{13} + 36687651\theta^{14} - 9532229\theta^{15} + 2421571\theta^{16} - 789888\theta^{17}\\ &+ 384252\theta^{18} - 36620\theta^{19} - 33000\theta^{20})\end{aligned}$$

$$\begin{aligned}A_7 = \frac{1}{1393459200}&(2596113838125 - 24170114626875\theta + 102231888554250\theta^2 - 260579462293500\theta^3\\ &+ 448535321698800\theta^4 - 555432907362200\theta^5 + 517094266858960\theta^6 - 375780963592520\theta^7\\ &+ 221036998846510\theta^8 - 108918282356690\theta^9 + 46304642945044\theta^{10} - 17387855835152\theta^{11}\\ &+ 5877848600212\theta^{12} - 1816618125644\theta^{13} + 520243088240\theta^{14} - 139724911768\theta^{15}\\ &+ 35649056429\theta^{16} - 8752126699\theta^{17} + 2123687858\theta^{18} - 512578900\theta^{19} + 159451124\theta^{20}\\ &- 76725412\theta^{21} + 7390832\theta^{22} + 6518976\theta^{23})\end{aligned}$$

$$\begin{aligned}A_8 = \frac{1}{2786918400}&(54927280170000 - 566627819428125\theta + 2682469240439625\theta^2 - 7736682221312625\theta^3\\ &+ 15238229702347575\theta^4 - 21815658385237150\theta^5 + 23657372605415500\theta^6 - 20072748261651120\theta^7\\ &+ 13725312699724360\theta^8 - 7784907830887710\theta^9 + 3764527022899598\theta^{10} - 1589720169572030\theta^{11}\\ &+ 597998092660338\theta^{12} - 203661747185296\theta^{13} + 63643856216676\theta^{14} - 18455949085100\theta^{15}\\ &+ 5015433541748\theta^{16} - 1288875268621\theta^{17} + 316012419057\theta^{18} - 74683452857\theta^{19}\\ &+ 17206812911\theta^{20} - 3959710266\theta^{21} + 917625928\theta^{22} - 274699196\theta^{23} + 130927516\theta^{24}\\ &- 12712784\theta^{25} - 10997952\theta^{26})\end{aligned}$$

$$\begin{aligned}A_9 = \frac{1}{367873228800}&(87498905321953125 - 990501375405898125\theta + 5187178735947573750\theta^2\\ &- 16693940022758286750\theta^3 + 37021429415280358125\theta^4 - 60197430084916965225\theta^5\\ &+ 74688975720920291500\theta^6 - 72835276973437848100\theta^7 + 57252153087954369930\theta^8\\ &- 37153888621452377830\theta^9 + 20381147835157710700\theta^{10} - 9665915767049820828\theta^{11}\\ &+ 4043679024080409186\theta^{12} - 1517882214148180122\theta^{13} + 518572009562370720\theta^{14}\\ &- 163168251434890064\theta^{15} + 47757902695290545\theta^{16} - 13114366898840337\theta^{17}\\ &+ 3404211749910806\theta^{18} - 841141673492318\theta^{19} + 199188193535793\theta^{20}\\ &- 45532047855261\theta^{21} + 10135109982756\theta^{22} - 2220037214076\theta^{23}\\ &+ 489502157056\theta^{24} - 109704211804\theta^{25} + 31805022744\theta^{26} - 15031625080\theta^{27}\\ &+ 1468402912\theta^{28} + 1248166272\theta^{29})\end{aligned}$$

$$\begin{aligned}A_{10} = \frac{1}{735746457600}&(2372826356485200000 - 29240694600135046875\theta + 167803910854979293125\theta^2\\ &- 596004415629137274375\theta^3 + 1469539466407887769125\theta^4 - 2676363324274425757125\theta^5\\ &+ 3744445956169018359875\theta^6 - 4139027045511795939425\theta^7 + 3697418806011442209775\theta^8\\ &- 2724220669452656462110\theta^9 + 1688684964920190890730\theta^{10} - 898195109034394895790\theta^{11}\\ &+ 417751651753405264878\theta^{12} - 172839123005311083618\theta^{13} + 64566413831667099198\theta^{14}\\ &- 22054839636256849378\theta^{15} + 6962544618355201150\theta^{16} - 2049840971510161799\theta^{17}\\ &+ 567169799383305641\theta^{18} - 148477226527006515\theta^{19} + 36996091151806065\theta^{20}\\ &- 8822528635761777\theta^{21} + 2024358244419159\theta^{22} - 449400090479373\theta^{23} + 97119395301731\theta^{24}\\ &- 20593013110736\theta^{25} + 4327327990744\theta^{26} - 920836701928\theta^{27} + 200624726700\theta^{28}\\ &- 56593505304\theta^{29} + 26545885944\theta^{30} - 2605385952\theta^{31} - 2179301760\theta^{32}).\end{aligned}$$

Appendix II: Polynomials $\mathcal{A}^{(j)}$ for $j = 1, \ldots, 5$

The generating functions $\sum_k A_k^{(j)} x^k$ providing the correction terms in formula (7) are given by the polynomials $\mathcal{A}^{(j)}(x, v)$ evaluated at

$$v = S(x) = -\frac{2}{x}\left(1 + (2x)^{1/3}\frac{\mathrm{Ai}'((2x)^{-2/3})}{\mathrm{Ai}((2x)^{-2/3})}\right).$$

$$
\begin{aligned}
\mathcal{A}^{(1)} &= \frac{1}{12}\left(2x + x(7x-2)v\right)\\
\mathcal{A}^{(2)} &= \frac{1}{1440}\left(5x(-2+7x)^2v^2 + (980x^2 + 80 - 376x)v - 80 + 456x\right)\\
\mathcal{A}^{(3)} &= \frac{-1}{25920x}\Big(5x^2(-2+7x)^3v^3 + 6(-2+7x)(245x^2 + 20 - 94x)v^2x +\\
&\quad (-320 + 16152x^3 - 6720x^2 + 1152x)v + 320 - 5864x^3 + 6832x^2 - 1312x\Big)\\
\mathcal{A}^{(4)} &= \frac{1}{43545600x^2}\Big(525x^3(-2+7x)^4v^4 + 840(245x^2 + 20 - 94x)(-2+7x)^2v^3x^2\\
&\quad + 56(2800 - 22640x + 462735x^4 - 336540x^3 + 109668x^2)v^2x\\
&\quad + (-62720x - 12969216x^3 + 27938400x^4 + 2989952x^2 + 89600 - 9632880x^5)v\\
&\quad + 8786688x^3 - 19265760x^4 + 17920x - 1811712x^2 - 89600\Big)\\
\mathcal{A}^{(5)} &= \frac{-1}{130636800x^3}\Big(105x^4(-2+7x)^5v^5 + 210(245x^2 + 20 - 94x)(-2+7x)^3v^4x^3\\
&\quad + 28(-2+7x)(321405x^4 - 237360x^3 + 82788x^2 - 16960x + 2000)v^3x^2\\
&\quad - 4(-232960x + 33600 + 20242740x^4 + 1524992x^2 - 7258616x^3 - 27277500x^5 + 4214385x^6)v^2x\\
&\quad + (10940288x^3 - 107520x + 94748736x^5 - 2745344x^2 - 42289728x^4 - 35840 - 67430160x^6)v\\
&\quad + 35840 + 1589504x^2 + 23671872x^4 - 6668160x^3 + 197120x - 56486304x^5\Big).
\end{aligned}
$$

Philippe Flajolet, Algorithms Project, INRIA Rocquencourt, 78153 Le Chesnay (France).
E-mail address: Philippe.Flajolet@inria.fr.

Bruno Salvy, Algorithms Project, INRIA Rocquencourt, 78153 Le Chesnay (France).
E-mail address: Bruno.Salvy@inria.fr.

Gilles Schaeffer, Loria, Campus Scientifique, BP 239, 54506 Vandœuvre-lès-Nancy (France).
E-mail address: Gilles.Schaeffer@loria.fr

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399